KI DELTA LEARNING
Scalable AI for Automated Driving

Projected GANs Converge Faster

Axel Sauer, Kashyap Chitta,
Andreas Geiger

WP 2.1
WP 2.4

Use-Case in Delta Learning

Generative Adversarial Networks (GANs) are widely used for domain transfer, e.g. Cycle-GAN, and data-augmentation with controllable image synthesis, e.g. Generative Radiance Fields (GRAF). This makes them a fundamental tool for closing the domain delta in existing datasets. However, GANs are challenging to train: They need careful regularization, vast amounts of compute, and expensive hyper-parameter sweeps. Our approach achieves significant headway on these issues improving image quality, sample efficiency, and convergence speed.

Technical Problem

In GAN training, the discriminator aims to distinguish real from fake samples. On closer inspection, the discriminator's task is two-fold: First, it projects the real and fake samples into a meaningful space, i.e., it learns a representation of the input space. Second, it discriminates based on this representation yielding the training signal for the generator. In our work, we explore the utility of **pre-trained representations** to facilitate the discriminator's task and improve and stabilize GAN training.

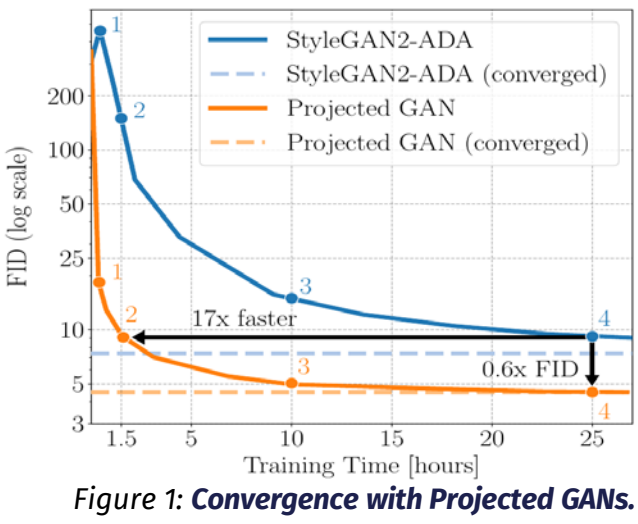

*Figure 1: **Convergence with Projected GANs.** Our approach significantly speeds up convergence while yielding lower FIDs.*

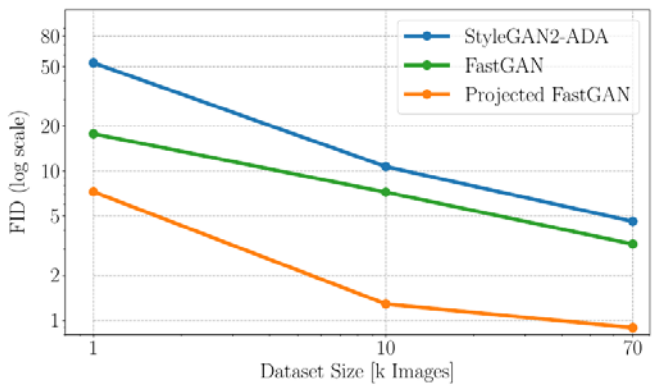

*Figure 4: **Small datasets**. Projected GANs significantly improve FID scores, even for subsets of CLEVR with 1k and 10k samples.*

Technical Solution

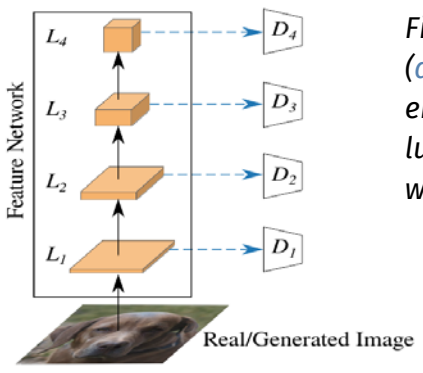

*Figure 2: **CCM** (dashed blue arrows) employs 1×1 convolutions with random weights.*

Cross-Channel Mixing (CCM). Empirically, we found two proper-
Ties of the projection to be desirable: (i) it should be information preserving (ii) it should not be trivially invertible The easiest way to achieve this is by mixing across channels with random 1×1 convolutions.

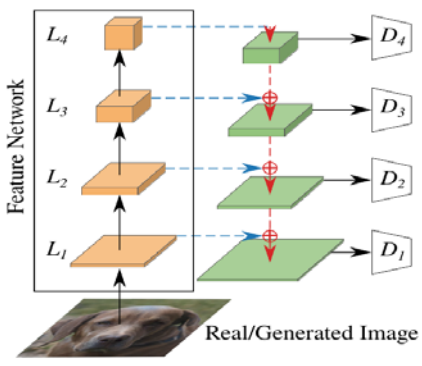

*Figure 3: **CSM** (dashed red arrows) adds random 3×3 convolutions and bilinear upsampling, yielding a U-Network.*

Cross-Scale Mixing (CSM).
To encourage feature mixing across scales, CSM extends CCM with random 3×3 convolutions and bilinear upsampling, yielding a U-Net architecture.

Evaluation

In our ablations on various pre-trained feature networks, we find that EfficientNets outperform both ResNets and Transformers. We thus use EfficientNet-Lite1 as our feature network. Our method achieves SOTA performance on both large and small scale datasets

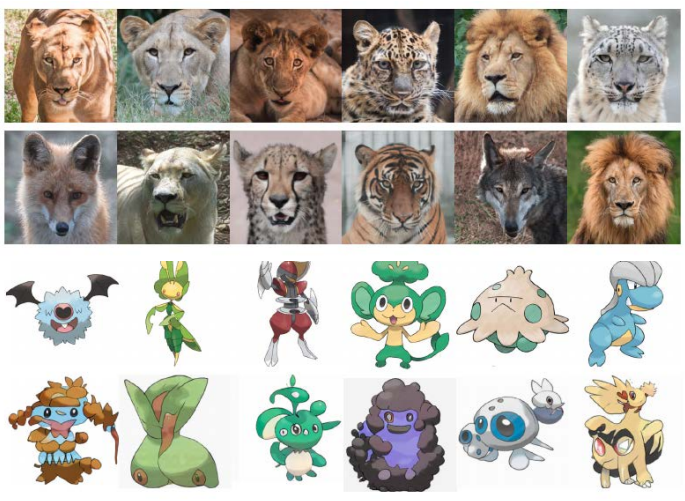

*Figure 5: **Real samples (top rows) vs. samples by Projected GAN (bottom rows)** for small datasets with 4812 and 833 images, respectively.*

Networks

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

For more information contact:
axel.sauer@tue.mpg.de

Partners

External partners

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Energy.

KI FAMILIE

Supported by: Federal Ministry for Economic Affairs and Energy on the basis of a decision by the German Bundestag

www.ki-deltalearning.de @KI_Familie KI Familie